DOCUMENT RESUME

ED 066 923 EM 010 319

AUTHOR	Campeau, Peggie L.
TITLE	Selective Review of the Results of Research on the Use of Audiovisual Media to Teach Adults.
INSTITUTION	Council of Europe, Strasbourg (France).
PUB DATE	Nov 71
NOTE	29p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Achievement; *Adult Education; *Audiovisual Aids; Cognitive Measurement; Comparative Analysis; *Instructional Media; Instructional Television; Learning Modalities; *Literature Reviews; Programed Instruction; *Research Methodology; Research Needs

ABSTRACT

The purpose of this literature review was to summarize results of experimental studies on the instructional effectiveness of audiovisual media in post-secondary education. Studies which met seven major screening criteria were used. A study was generally accepted if it compared performance of experimental and control groups on objective measures of cognitive achievement. Of 200 items studied in detail, only about a dozen met screening criteria. These are summarized under the following headings: programed instruction (the largest group); television, videotape, and motion pictures; slides, filmstrips, overhead transparencies, and still pictures; radio and tape recordings; and multiple media. A closing section enumerates common methodological defects in recent media research. It is proposed that potentially interacting variables be 1) defined according to some theoretical framework, 2) used in multivariate designs, and 3) taken into account in planning more multi-media studies. More effort should also be devoted to replication and follow-up studies. (Author/JK)

ED 066923

SELECTIVE REVIEW OF THE RESULTS OF RESEARCH ON THE USE OF AUDIOVISUAL MEDIA TO TEACH ADULTS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Prepared for the Council of Europe

by

Peggie L. Campeau

November 1971

EM 010 319

SELECTIVE REVIEW OF THE RESULTS OF RESEARCH ON THE USE OF AUDIOVISUAL MEDIA TO TEACH ADULTS

Introduction

Purpose of the Review

The purpose of this literature review is to summarise findings from experimental studies which have assessed the instructional effectiveness of various audiovisual media for teaching adults. The studies cited here were selected in the light of the objectives of a larger project, of which this review is one preliminary phase. As discussed by the Council of Europe, whose interest is largely in the field of adult education, the larger project will be an attempt to construct a media taxonomy which can be used by teachers to select appropriate media for specific learning tasks. To construct the taxonomy, decisions will have to be made about which media are most appropriate for each of the specific learning tasks represented in the taxonomy. In anticipation that there might be some experimental research evidence to help make these decisions, the present literature review was carried out.

As will be made clear in the following section, the needed evidence was sought in the "hard" research data provided by experimental studies of media effectiveness. This literature contrasted with the "softer" insights into instructional media that were provided in the nonexperimental literature, most of which did not attempt to deal with or assess instructional effects.

Scope of the Literature Review

This review was strictly delimited by the objectives envisioned for the proposed taxonomy. Comparative effectiveness studies and research on media, learner and task variables were judged to be of the most potential value for helping to develop a taxonomy teachers could use to select appropriate media for specific learning tasks. The literature review has focused on these types of research, providing that such studies appeared to meet several screening criteria. These criteria indicate the measures taken to ensure that a study's findings were interpretable.

Seven major criteria were used to select studies for this review. Three criteria were related to the nature and quality of reported information, while the other four pertained to experimental design.

Recency. This review deals only with evaluation data published during the six-year period from 1966 through 1971. This criterion was imposed to avoid repeating work done during a previous review of research included in the text on instructional media by Briggs, Campeau, Gagne and May (1966). The chapter by Campeau, summarising research data published

as late as 1965, reported findings for each media category by age group and may be referred to as background for the present review.

Completeness. The only studies included in this review are those in which there was reported enough detail on experimental procedures to permit a study's conclusions to be evaluated. In no case was an abstract accepted as the sole source of information about a study.

Availability. No unpublished studies were considered. Every study cited is available in the form of a journal article, as a published book or document or on microfilm or microfiche.

Sample. This review is limited to studies in which adult (post-secondary age) learners were assigned to experimental and control groups that included, normally, at least 25 subjects each. The restriction to studies using adult learners was made because the proposed media taxonomy will be designed primarily for use in post-secondary education. In addition, the groups being compared must have been properly comparable. Reasonable proof of similarity was presumed to be evidenced by random assignment of subjects to treatments or by matching subjects on the basis of relevant variables which, if uncontrolled, could account for effects attributed to the mode of instruction. At the very least, there had to be some reported basis for assuming similarity, e.g., similar pretreatment scores on achievement, ability or content measures.

Treatment. For a study to be included in this review, the treatment must have been administered for no less than one hour and have been unconfounded by differences in content coverage among groups being compared.[1] Again, because the proposed media taxonomy will be designed primarily for use in post-secondary education, all studies which did not concern the teaching of concepts or skills similar to those in lessons or lectures for adults were excluded. Much of the research reported in psychological journals was omitted because the treatment consisted of highly specific laboratory-type tasks which had no obvious, direct and practical application to instructional practice or to the type of media taxonomy planned for the Council of Europe. The requirement that treatment conditions exhibit equal content coverage was easy to apply because investigators nearly always specified the extent to which content coverage was similar for all comparison groups. On the other hand, it was very difficult to evaluate the quality of experimental designs, and biases introduced by uncontrolled variables may still be concealed in the studies cited here.

Measures of media effectiveness. Studies summarised in this review are those in which media effectiveness was assessed by scores on objective achievement tests (as opposed to course grades), and in which identical tests were administered to groups being compared. It was not possible to determine in all cases whether the instruments used were valid or reliable, because these data were rarely reported. Neither was it possible to draw conclusions as to whether statistically significant gains, where reported,

---

[1] This criterion would have been waived, had studies been found in which differential content coverage by the media being compared was correlated with performance on different types of test items. No such studies were found which also met the other screening criteria.

were educationally significant as well. For example, educational significance might be shown by a rate of gain from pre- to posttest which exceeded that expected on the basis of test norms. Or, one might use a rule of thumb that an educationally important gain is one which exceeds one-third of a standard deviation. The assumption here is that gains of less than that amount have little educational significance. Neither reference point could be applied, however, because standardised criterion tests were seldom used and standard deviations were rarely reported.

Statistical reliability. To be selected for review, reports of studies must have described the statistical techniques that were used to analyse data, and must have reported differences between groups as being either statistically significant or nonsignificant. In the case of a significant difference, it must have been evident that the difference could have occurred by chance no more than five times in one hundred replications ($p < .05$).

Other Exclusions

Utilisation studies. Research on how a given medium teaches best was not included in this review. Such studies usually compared one method of using a medium with another method of using the same medium to teach the same instructional content. Instead, this review is restricted to comparisons of the relative amounts of learning which result from using different instructional media to teach adults. Also included are the more analytical studies which explore interactions among media, learner and task variables as they relate to achievement.

CAI studies. Research on computer-assisted instruction was not considered at all, on the assumption that such systems are still too costly for widespread use in post-secondary education. Evidence from CAI research would therefore be unproductive and impractical for the present project, given that the proposed taxonomy must be useful to teachers of adults.

Other categories of research omitted from review were historical or descriptive studies, reports of surveys, attitudinal and motivational research and media-preference research.

Search Procedures

The literature search was conducted to identify books, documents and journal articles relevant to the audiovisual instruction of adults. Procedural details are presented here, not so much to attest to the thoroughness of the search, but more importantly to emphasise the appalling lack of recent, objective data on the instructional effectiveness of audiovisual media, particularly in the field of adult education.

The literature search utilised both computerised and manual procedures as enumerated below:

Educational Resources Information Center (ERIC). The major source of published research reports and journal articles was the computer-based ERIC collection. Supported by the U.S. Office of Education, ERIC is a national information network for acquiring, abstracting, indexing, storing,

retrieving and disseminating significant and timely educational research reports originating in and outside of the United States. Also included are studies in which military, civil service, industrial and technical personnel served as subjects.

The search of this collection was made by computer. At the time the search was run, the ERIC data base contained about 100,000 items, 1200 of which were identified as being more or less relevant to the use of media to teach adults. It was not possible to eliminate nonresearch studies without studying the printout. The computer printout supplied abstracts for nearly all of the 1200 entries. These were read to decide if microfiche of the original documents should be studied and if journal articles should be read. Where no abstract was provided, as in the case of most journal citations, such decisions had to be based on information contained in the title and in the descriptors which had been assigned to the item by ERIC personnel. Admittedly, documents whose value was not conveyed by title, descriptor or abstract may have been overlooked.

These procedures resulted in the identification of 97 reports and 92 journal articles for in-depth review. An estimated three-fourths of these documents and journal articles concerned media research or research-related activities conducted within the United States. The remaining one-fourth pertained to similar activities outside the United States, notably in Great Britain Canada and France.

The 97 documents were read on microfiche borrowed from the ERIC Clearinghouse on Educational Media and Technology at Stanford University. The 92 journal articles were read by obtaining appropriate issues of journals from the various technical libraries in or near the university.

British collections. An especially productive source relating to media research in Great Britain was the four-volume series edited by Coppen (1968, 1969, 1970, 1971) containing abstracts of British research in audiovisual aids from 1945 to 1970. Original documents or journal articles for many relevant studies were obtained for in-depth review.

Additional manual search. Materials identified through the computer search yielded research reported as late as 1970. Insofar as time permitted, further updating was accomplished through a manual search of current selected British and American journals devoted to educational research, and of late-1971 issues of Education Index (U.S.), the British Education Index, Current Index to Journals in Education and Research in Education. (The last two collections are ERIC publications.)

Technical libraries. The book collections of the ERIC Clearinghouse on Educational Media and Technology and the Palo Alto offices of the American Institutes for Research were surveyed for additional sources of media research.

Mail inquiries. In addition to the computerised and manual searches described above, mail inquiries were sent to six U.S. experts in the field of media research. These individuals were asked to inform the author of any current media research that might fit the screening criteria for this review, but for which findings might have been published too recently to be included in the ERIC collections or in the current periodical indexes. Four of these individuals responded in time for their suggestions to be followed up.

For all practical purposes, the search for media studies was concluded during October, 1971, when writing of the review was begun.

Disappointing Yield from the Literature Search

Considering the relatively small number of studies cited in the following pages, it is evident that most of the current or recent media research does not meet the criteria established for this review. Moreover, reports of research constituted only a small fraction of the vast body of literature on the use of instructional media to teach adults that was published during the period covered by this review. Most of this nonexperimental literature consisted of surveys, testimonials, historical and descriptive assessments, reports of informal evaluations--all of which did not even attempt to deal with or assess the instructional effectiveness of audiovisual media. Instead, assessments were made of user preference for and attitudes toward various media, patterns of media use, characteristics of the post-secondary audience, problems of educating adults via mass media and so on. Much of the data offered were in the form of questionnaire responses, enrollment and completion statistics, cost figures, and tallies of services provided and extent of use.

Particularly disappointing was the low yield of suitable studies from military and industrial training research. Such studies typically did not meet the screening criteria, either because treatment duration was too brief, sample size was too small, instructional effectiveness was not measured in terms of cognitive benefits, or original documents were not available within budget, time and accessibility limits of this project.

The Literature Review

The research summarised below pertains primarily to the use of programmed instruction or television to teach adults. Other categories of research deal with instruction of adults by videotape, motion pictures, still pictures, radio, tape recordings and various combinations of media. The inclusion of studies under each media heading was determined according to the presentation device. Thus, a televised presentation in which the lecturer plays tape recordings and shows a film would be reported under the heading, Television.

Within each media category, reviews of research are summarised first, followed by descriptions of individual studies which were found to meet the criteria for this review.

Programmed Instruction

By far the largest category of media research reported from 1966 to 1971 was programmed instruction. In most of the studies which met the criteria for this review, the instructional effectiveness of programmed materials was compared with that of face-to-face instruction by a teacher. The remaining few studies which also met the criteria investigated the interactive effects of learner, task and program variables on achievement. Before presenting findings from individual studies, a few significant

reviews of programmed instruction research are cited. It should be borne in mind that the reviewers seldom provided enough information about individual studies to permit our screening criteria to be applied. Occasionally, reviewers mentioned that no stringent selection criteria had been used, that abstracts were accepted in lieu of original documents and that their summaries amounted to collations of research. Nevertheless, the conclusions drawn by these reviewers provide the reader with a frame of reference for interpreting the findings of the studies cited here.

Recent reviews of programmed instruction research. Studies comparing programmed with conventional instruction have yielded mixed results across a wide range of subject matters and age groups (Campeau, 1966).

McKeachie (1967) reviewed media research in higher education. He indicated that comparative effectiveness studies at the college and university level have produced inconsistent results when face-to-face instruction was compared with instruction by Skinnerian programs.

Briggs (1968, 1970) reviewed studies in which learner variables were related to program variables. No clear pattern of interactions emerged between personality, intellectual or attitudinal characteristics of learners and program variables such as pacing, branching and size of step. Briggs suggested that disparate results can arise from the use of a wide variety of learning tasks, variations in programming skill, differences among subjects in prior learning or lack of precise definition of variables under study. Further, the studies reviewed by Briggs encompassed a wide age range. It may be possible that age-related learner variables interacted in different ways with the program and task characteristics investigated in these studies.

Research on student response aspects of programmed instruction has been somewhat more fruitful. In reviews of recent studies on the role of response mode in programmed instruction (Allen, 1971; Anderson, 1967; Campeau, 1966; McKeachie, 1967; Tobias, 1968, 1969; Tobias and Abramson, 1970), it was concluded that there was no advantage in requiring learners to make overt responses unless the responses were difficult or novel (technical terminology), or unless the responses themselves had to be learned (typewriting, vocabulary).

Finally, several reviews of research on the structure and sequencing of instruction were found in which particular attention was devoted to task analysis and presentation of skills to be taught (Briggs, 1968; Gagne and Rohwer, 1969; Frase, 1970; Rothkopf, 1970). However, in much of the research cited, programmed instruction was not the presentation device and many participants were elementary or secondary school students. Also, the conclusions drawn were not so much conclusive as suggestive of future research.

Comparative effectiveness studies. Eight studies which appeared to meet the screening criteria for this review investigated the instructional effectiveness of programmed materials when compared with the traditional means of instruction in the course, whether by lecture, lecture and discussion, lecture and textbook, lecture and laboratory or some other combination.

Of these eight comparative effectiveness studies, five found no significant differences which could be attributed to instructional method; they are summarised first. The remaining three studies reported significant differences in favour of programmed instruction.

Wilds and Zachert (1966) conducted a two-year investigation of the use of programmed instruction in gynecologic oncology. Six medical schools were involved in the seven studies they report. Of the seven studies, the one conducted at the University of Georgia during the first year of research appeared to meet all the criteria for this review, and the following summary is restricted to that eight-week study. The entire junior class (N=90) was divided into experimental and control groups by stratified random assignment based on grade point average during the first two years in medical school. The experimental group used an 830-frame, linear, "content" program designed to replace eight hours of conventional classroom teaching of gynecologic oncology. The control group attended eight one-hour class sessions where they listened to tape-recorded lectures which were specially prepared to duplicate the coverage provided by the programmed instruction. These lectures were supplemented by the extensive use of slides and other visual aids. The relative effectiveness of the programmed text and lecture presentations was evaluated on the basis of gains from pre- to posttest administrations of alternate forms of a standardised test on gynecologic neoplasms. Analysis of gain scores indicated that there was no significant difference between treatments.

In another setting, Barnes (1970) explored the effectiveness of programmed instruction in a university physical science laboratory. The programmed materials were used only during laboratory investigations, in contrast to most of the other studies cited here in which programs were used as the sole or major method of instruction. Since scheduling difficulties precluded random assignment of individuals to treatment groups, Barnes designated six lab sections as experimental (N=105) and six lab sections as control (N=109), based on lecturer, time of day and time of week. Regardless of designation, students attended lectures as usual, covered similar subject matter and were assigned the same series of laboratory investigations. Four out of seven of these laboratory investigations are the basis for the procedures and results summarised here. Students in the experimental lab sections used specially prepared programmed materials which contained introductory information, directions for laboratory procedures, blank data tables and questions to be answered. No further assistance or instruction was provided. The control students used the conventional laboratory manual. This manual was narrative in format and contained introductory textual material, directions for laboratory procedures and questions for students to answer. In addition, the laboratory instructor lectured for several minutes at the start of each period to explain further the principles and procedures with which the students would be involved in their investigations. On pretests measuring scholastic aptitude and knowledge of physical science subject matter, no significant differences were noted between experimental and control students. The latter test was readministered as a posttest, along with another specially prepared laboratory performance test. No significant differences between groups were found on either of these post measures. In subgroup analyses, no significant interactions between treatment and ability were noted.

A study by Unwin (1966) compared conventional and programmed instruction in terms of short- and long-term test performance. The entire enrollment (N=45) for a first-year undergraduate engineering course was assigned to two treatment groups on an alphabetical basis. One group worked through a 100-frame programmed text during the ten hours of class time in which the conventional group received regular lectures and tutorials

on the same mathematics subject matter. An initial familiarity with elementary algebra was assumed, and instruction was aimed at training students to evaluate determinants and to use them in solving simultaneous and homogeneous equations. At the end of the two-week experiment, a posttest was administered and scored by a lecturer who did not know about the experiment. No significant differences between groups were found. Retention test data were collected eight months after treatment, but are not reported here. The loss of several subjects reduced the size of each group well below the minimum set by the screening criteria for this review.

Giese and Stockdale (1966) found no significant differences which could be attributed to programmed or conventional workbook methods of teaching college students grammar, sentence structure, punctuation and capitalisation. The basis for measuring instructional effectiveness was pre- and posttest performance on standardised English tests. The students using the program were reported to have mastered the material in a shorter period of time than students in the conventional group.

Alexander (1970) also failed to find a significant difference favouring the use of either a dictionary or a program to promote vocabulary growth. In his study, 90 freshmen enrolled in compulsory remedial reading classes at a Kentucky senior college were assigned by intact classes to experimental and control groups. According to pretest data, the groups were comparable in mean level of abstract intelligence. During the ten-week study, two teachers worked with students, nine students at a time, providing similar amounts of vocabulary instruction to students in both experimental and control groups. As an adjunct to instruction in these small groups, control students used a dictionary to study word meanings, completed an exercise sheet using words studied and wrote compositions or sentences using as many of the words studied as possible. The experimental students completed the same composition and sentence-writing assignments, but used published programmed materials instead of the dictionary to supplement instruction by the teacher. On the basis of statistical analysis of pre- and posttest scores on a standardised vocabulary development test, the investigators found no significant differences between the two methods of promoting vocabulary growth, regardless of sex or level of abstract intelligence.

In contrast to the inconclusive findings from these five investigations, three studies did report significant differences in favour of programmed instruction. Marsh and Pierce-Jones (1968) assessed the effects on learning of adjunct programmed materials, various types of presentation devices, self-paced and imposed laboratory schedules and an initial cognitive organiser in the form of a multiple-choice pretest. The entire enrollment (N=295) in all eight sections of a semester-long course in adolescent psychology participated in the study. Students were randomly assigned to the various cells of a complex 4 x 2 x 2 x 2 balanced factorial analysis of variance design. All combinations of the experimental conditions were present in each of the eight sections in order to control for teacher differences. It was determined statistically that there were no initial differences between experimental and control groups, or among students assigned to the various instructors, on scholastic aptitude, reading achievement and English achievement test scores. At the end of the semester, a 100-item multiple-choice test was given to all students. Based on analyses of these scores, the investigators drew several conclusions. Most pertinent

FILMED FROM BEST AVAILABLE COPY

to this review was the finding that students who worked through programmed materials as an adjunct to regular instruction scored significantly higher on the test than students who attended the same laboratory, but completed tasks other than programmed instruction. Another finding was that no significant differences in learning resulted from presenting the programmed material by teaching machine, as opposed to presentation by programmed workbook. Similarly, no significant differences in criterion test scores were noted for students whose laboratory schedule was determined for them, and for those who arranged their own schedules for completing laboratory assignments. On the other hand, the mean test score of those students who took the same criterion test at the beginning of the semester (as an advance cognitive organiser) was significantly higher than the mean score of those students who were not given such a pretest.

Hamilton and Heinsohn (1967) studied the efficacy of a programmed text to teach "review English." Matched classes of 58 students were formed on the basis of scores on a standardized English screening test. In one group, the main source of instruction was a published programmed text; the other group used the conventional text. The experimental group performed significantly better than the control group on an examination designed to assess learning from the programmed material. On another examination, more appropriate for the control group, no significant differences were noted. Both examinations dealt with the broad objectives of the course (grammar and punctuation), but only the program-relevant test correlated highly with grades earned in the course by both groups. Apparently the program better covered that material which students had to learn in order to achieve course objectives.

In a study conducted by Daniel and Murdoch (1968), use of a programmed textbook was compared with use of a conventional textbook to see which method resulted in better performance on a content examination for a college course in operant psychology. In this experiment, 12 teaching assistants each taught two discussion sections of about 22 college undergraduates each. One of the sections studied the programmed text and the other studied the conventional text. Both texts covered similar material. At the end of the semester, all students took a 100-item exam containing several types of items. Results of a multivariate analysis of variance, used to assess the effects of the experimental variables on these several types of items, indicated that for all six objective types of items and for five out of eight essay items the level of performance of the programmed text group was significantly better than that of the conventional text group.

The focus of the eight studies just summarised was mainly upon the overall effectiveness of programmed instruction when it replaced part or all of a more traditional, established or conventional mode of instruction. No particular attention was devoted to isolating or manipulating program or learner characteristics. In the three studies which follow, some attempt was made to take program, task and learner characteristics into account, whether in designing the experiment, analysing the data or interpreting the findings.

Studies of program, task or learner variables. Hartley (1968) compared the instructional effectiveness of a branching program with that of the regular textbook in a college course in statistical methods. He also sought

to identify noncognitive traits which helped determine whether or not a particular student would benefit from programmed instruction. One hundred thirty-eight university students were divided into two groups equated on sex and mathematics background and assigned either to programmed text or regular text treatments. For four weeks, the first group (N=38) worked through a branching program on descriptive statistics. Program frames were presented by machine and included frequent requirements for written responses, provision of feedback and completion of review tests and examples. The other group (N=100) received conventional instruction which was characterised as unstructured and consisted of assigned reading which covered equivalent material in the regular course text. No pressure was put on either group to finish the instructional material, other than stating that an exam would be given at the end of the four-week period. Students in both groups were given the same posttest, in addition to completing questionnaires and scales designed to assess note-taking habits, attitudes and personality traits. Results of the posttest and personality scales are of most interest in the light of objectives for this review. The posttest, composed of items selected by teachers not involved in the study, consisted of content, application and transfer questions. Analysis of posttest scores revealed a significant difference in favour of the programmed instruction group. Moreover, this superiority held for performance on all three types of test items. Analyses of scores of "finishers" and "non-finishers" indicated that the superiority of the program treatment was accounted for by those students who finished at least three-fourths of the programmed instructional material. Based on the fact that gains were differentially distributed between those who finished the program and those who did not, Hartley concluded that the program benefited the students who completed most of it, but had no advantage for the non-finisher and was more expensive than the textbook. The personality scales, which had been demonstrated in earlier research to correlate significantly with final statistics examination grades, consisted of items designed to measure traits of neuroticism and extroversion. For the finishers and non-finishers in Hartley's study, scores on the personality scales predicted which students were likely to complete the work required by the programmed text. Nearly all students whose scale scores identified them as "stable extroverts" were non-finishers. On the other hand, nearly all the "anxious introverts" were finishers. These observed frequencies were statistically significant. If the assignment of subjects to treatment groups had taken into account scores on the personality scales, one wonders what results would have been obtained. For example, finishers accounted for the superiority of the program group, yet only 53 percent of the program group were finishers. If all the "anxious extroverts" had been assigned to the program group, would the overall performance of this group have been considerably higher? It would seem so, by virtue of dramatically increasing the number of students who were likely to complete, and thus benefit from, the programmed text.

Karis, Kent and Gilbert (1970) employed a 384-frame technical program dealing with medical subject matter to investigate the role of responses per frame, response mode and response confirmation in linear self-instructional programs. The program was completed during four periods, one period each day for four consecutive days, by 450 university freshmen who were randomly assigned to one of 18 groups. Each group used a different version of the program. Sixteen versions were used to compare the main and interactive effects of overt and covert responding, confirmation vs.

nonconfirmation, limited vs. expanded frame content and single vs. multiple frame responses in a 2 x 2 x 2 x 2 factorial design. Two additional program versions (reading only) were used to provide control data. Criterion measures consisted of four daily unit tests, a comprehensive posttest and a delayed retention test administered about one month after termination of treatment. Items in each test required either the recall and reproduction of medical terms or the definition of medical terms. Findings were that requiring overt responses during learning of this material resulted in significantly better test performance than that noted for subjects who learned the same material under covert response conditions. It was also found that, although the overt-response groups took longer to complete the programmed material, they spent less time than the covert-response groups completing the tests. On all but the delayed retention test, the overt-response groups were also faster than the reading-only groups. Additional findings from this study were that withholding confirmation produced no differential effect for any group and that variations in frame size resulted in no significant difference. On the other hand, the variation of number of responses required per frame produced a significant effect in favour of multiple responses when test items required medical term responses.

Finally, the influence on student achievement of redundancy in programmed instruction was investigated by Valverde and Morgan (1970). In their study, students at an American Air Force medical service school were randomly assigned to five groups, each numbering 88. A programmed course teaching the prefixes, roots and suffixes of medical terminology was modified to produce three different versions, represented respectively by a 274-frame program, the same program reduced to 160 frames and reduced again to 83 frames. Two additional versions of the same material were a typographically cued response mode (with important words underlined) and a summary card on which all the terminology information was presented. Pretest data from a comparable uninstructed group had indicated that the typical student had little, if any, prior knowledge of the subject matter. On the multiple-choice posttest, similar achievement was attained by groups using the minimally redundant (83-frame) program, the brief narrative (typographically cued) coverage of the same information and the study card which presented all essential terminal behaviours. These three groups were superior to the highly and moderately redundant program versions. The authors concluded that eliminating the redundancy in the linear program significantly increased student achievement on the medical terminology posttest. Whether these results would have been obtained on a constructed-response (recall) test, as opposed to the multiple-choice (recognition) test, is open to question.

Television, Videotape and Motion Pictures

Only three studies were found which met all the criteria established for this review, and all three investigated the instructional effectiveness of television. With the exception of these studies, most of the following discussion summarises other summaries of research on the instructional effectiveness of television, videotape and motion pictures--hardly a good basis upon which to build a media-selection strategy. Further, the conclusions reached by reviewers were based mainly on research at elementary and secondary grade levels. The degree to which these conclusions might hold for instruction at the post-secondary level has not been adequately tested by recent media research.

Television. Major reviews of literally hundreds of comparative effectiveness studies concluded that, in general, no significant differences were found when instructional television was compared with face-to-face, live instruction (Allen, 1971; Campeau, 1966; Chu and Schramm, 1967; Dubin, Hedley, Schmidbauer, Goldman and Traveggia, 1969; McKeachie, 1967; Reid and MacLennan, 1967; Twyford, 1969). Dubin and others conducted an intriguing analytical review of studies in which experimental classes taught by television were compared with control classes receiving no televised instruction. The authors identified 42 such studies that could be considered to be comparable on the basis of several criteria, namely, instruction lasted at least one semester; identical, written, course-content examinations were used for groups being compared; similar methods of instruction were experienced by both groups, whether in a televised or face-to-face situation, and so on. Before combining data from the 42 studies meeting the selection criteria, the authors calculated t-ratios for each of nearly 100 independent comparisons as a standardised measure of the amount and direction of difference in group means. Based on statistical tests of t-ratio distributions, the authors drew several conclusions. First, when teaching methods were matched, face-to-face instruction was only superior to two-way instructional television, and then only when the lecture method was used by each medium. ("Two-way" television provided students and lecturer with audio facilities for exchanging questions and initiating discussions, thereby approximating a "live" instructional situation.) Second, one-way instructional television produced the same amount of learning as face-to-face teaching by lecture, a combination of lecture-discussion-demonstration or discussion alone. Third, instruction by either method yielded no significant differences when the studies were grouped by the broad subject-area headings of humanities, social sciences and science-math. In attempting to explain the clear finding that two-way television was definitely inferior to face-to-face teaching (both using lecture methods), the authors conjectured that the requirement for students and lecturer to utilise the fairly complicated technical apparatus necessary for two-way communication may have been detrimental to the effectiveness of the medium.

Chu and Schramm's review was based on 207 studies involving 421 separate comparisons. The investigators stated that consistencies across separate comparisons indicated that instructional television was less effective at the college level than at the high school or grade school levels. At the college and adult level, results of 235 comparisons indicated that 176 found no significant differences between televised and conventional instruction, 29 favoured television over conventional instruction and 30 comparisons were significant in the opposite direction. In those studies using adult participants, there seemed to be no conclusive or consistent evidence to suggest that the following variations would improve learning from television: physical variations such as size of screen, use of colour, camera angle; variations in viewing conditions pertaining to viewing angle and distance, home or school viewing, homogeneity or heterogeneity of viewing groups, permissive versus required viewing; pedagogical variations such as use of humor and animation, dramatic versus expository presentation, use of inserted questions; variations in student response mode; variations in student-teacher contact, such as two-way talkback. The latter finding is at odds with the research reviewed by Dubin and others in the preceding paragraph. Chu and Schramm did find some evidence that feedback as to test results had a positive effect on learning from instructional television.

McKeachie (1967) cited his own and other research in support of the conclusion that television is less effective than "live" instruction for college and university students. He stated that although differences in individual studies were not statistically significant by themselves, "Their consistency was statistically significant." He cautioned that to be properly evaluated, the effectiveness of televised instruction should be measured by appropriate criterion instruments. According to McKeachie, achievement tests now used do not measure the proficiency with which students can evaluate visual properties of the instructional content, and these properties might be taught better by televised instruction.

The three experimental studies which met the criteria for the present review are summarised below. One study compared televised with non-televised instruction in an engineering course, one compared televised with tape-recorded lectures and one investigated the contribution of colour to learning from television.

Smith (1968a, 1968b) conducted a series of experiments to obtain evidence of the effect of television broadcasts on the attainment and attitudes of students taking "G" courses at technical colleges in England. The third and largest experiment, summarised here, involved students in 27 technical colleges where the BBC television series on engineering science was broadcast. At most of the colleges, all participants watched the television series; however, eight of the colleges contributed students to both experimental and control groups. Of the 862 students with complete sets of data on measures of ability, attainment and attitude, 382 had seen the television broadcasts and the rest constituted the control group. Analysis of pooled test data from all 27 colleges indicated that the two groups were well matched on measures of verbal and nonverbal mental ability. However, there also were no significant differences between groups on the attainment test used to assess the instructional effects of the two treatments. On the other hand, large and significant differences both in ability and attainment were found when the scores for individual colleges were analysed separately. Reasoning that real differences between treatments could have been masked by factors such as variations in teaching efficiency, availability of equipment and library facilities, the investigator conducted an analysis of combined data from the eight colleges contributing to both experimental and control groups. It was evident from these analyses that the two groups, when selected in this way, were again closely comparable with respect to verbal and nonverbal mental ability, but were significantly different on measures of attainment and attitude. Students who had seen the television broadcasts did better on the whole on the engineering science test and had more favourable attitudes toward the course than those who had not seen the broadcasts. Correlational analyses were then performed on the data. Findings were that the television broadcasts appeared to have a greater impact on students who were above average in ability and to have more effect on performance in mechanics. The investigator suggested that visual presentation of subject matter might be especially helpful in teaching a subject requiring spatial thinking.

Rich, Poll and Williams (1966) reported two experiments, both using subjects from a required, college-level history course. One experiment compared instructional television with face-to-face instruction. In the second study, they compared large-screen televised presentations with those presented on small screens. Students drawn randomly from the entire course

were assigned either to one of two morning groups (televised versus face-to-face instruction) or to one of two afternoon groups (large-screen versus small-screen televised instruction). In the morning (first study), the experimental group (N=474) viewed televised lectures while the conventional group (N=127) met with the lecturer and listened to tape-recorded lectures. On an optional basis, students in both groups participated in weekly discussion sessions led by a graduate assistant. Occasionally, they also joined to participate in videotaped debates or to view films. The afternoon (second study) experimental group (N=385) viewed large-screen televised lectures which could make use of long shots, closeups, superimposed matter and other features not feasible in traditional televised instruction. The conventional group (N=47) viewed the lectures on two 24-inch television monitors. Again, the optional discussion sessions were available to both groups. The same evaluation procedures were used in the two experiments. Seven bi-weekly quizzes (a total of 30 objective questions) and a final exam (100 objective questions) were given to the experimental and control groups and were then machine scored. Only results of the final exams are considered here because quizzes averaged no more than four or five items each. In the first experiment, significant differences were noted between students who received lecture presentations on pre-recorded tape and students who viewed televised lectures. In the second experiment, in which all students received televised instruction, there were no significant differences between the large- and small-screen groups. Other factors which could have accounted for obtained results were also investigated: viewing angle and distance from screen. Results of statistical analyses showed no significant differences in performance. It was therefore concluded that there were no measurable differences in learning or information gain as a result of presentation device, size of screen or student location in the auditorium.

Kanner (1968) reported data from two television studies he conducted to investigate the contribution of colour to learning by American Army trainees. His studies compared the teaching effectiveness of colour and black and white television presentations. Analysis was also made of the effects of trainee aptitude and type of subject matter upon learning. The same televised instruction was presented simultaneously to two groups equated for aptitude. One group saw colour, while the other group saw black and white presentations. Immediate posttest results from both studies found no significant differences between treatments in 10 out of 11 comparisons over 11 different subject matters using various amounts of colour in charts, maps and other training materials. On the posttest, some multiple-choice items incorporated colour to help the trainee understand or answer the questions. When performance on these test items was analysed, no significant differences were noted between the two treatments. Kanner concluded that colour did not appear to provide important cues to learning not found in black and white presentations, perhaps because of verbal cues in the presentations. His findings are in accord with research reviewed by Chu and Schramm (1967).

Videotape. To date, media research has not dealt in any systematic way with cognitive achievement benefits which might accrue from videotaped instruction. Although videotape has been used widely to train prospective teachers in competencies basic to their teaching functions (Lesser and Schueler, 1966), assessment has been limited mainly to subjective judgments based on direct observation of teaching behaviour in the classroom. Thus,

no studies or reviews of research on the instructional effectiveness of videotape were found which fell within the purpose and scope of this literature review.

Motion pictures. In the search for useful film research too, no experimental studies were found which met all screening criteria. Conclusions about the instructional effectiveness of film are therefore limited to reviews of research.

Motion pictures and television have been considered to be quite similar presentation devices. Chu and Schramm (1967) concluded that there appeared to be little difference between learning from television and learning from film, if the two media were used in the same way. This may provide some explanation as to why Reid and MacLennan (1967) found that instructional film research exhibited the same general no-significant-difference pattern as that found for most comparative effectiveness research in instructional television.

It may be of interest to note parenthetically that much of the film research cited in the Reid and MacLennan review was at least ten years old, that relatively few comparative studies pertaining to the instructional effectiveness of film had been conducted in the fifteen-year period (1950-1964) covered by their review and that very often abstracts had to be relied upon because original documents were not available.

McKeachie (1967) reviewed film research (much of it carried out by the American armed forces) and summarised certain "emerging principles." It appeared that students learned at least as much from film as from a poor teacher. Participation was found to increase learning and McKeachie suggested it be planned in the production of a film or television presentation. Further, students having previous experience with instructional films learned more than students without such experience, at least for learners with little prior knowledge of the subject matter.

Slides, Filmstrips, Overhead Transparencies and Still Pictures

No suitable studies were found in which the instructional effectiveness of any one of these media was assessed. Two studies, in which slides and filmstrips were used in combination with other media, are reported under Multiple Media.

Chu and Schramm (1967) drew some conclusions about the relative effectiveness of still and motion pictures, based on their review of research. There was no consistent evidence that motion pictures enhanced learning more than still visual images. On the other hand, in two studies using college or military participants, some superiority for moving over still visual images was noted for tasks (or subtasks) in which continuity of the filmed action played an essential part in learning to perform the task.

Without drawing a distinction between still and motion pictures, Chu and Schramm suggested that beneficial effects of visuals on learning depended on the nature of the task, familiarity with the concepts being taught and test format. They speculated that visuals would enhance learning of a manipulative task and would assist the association of two concepts (provided

that one of the two to be associated was familiar, as in learning a foreign vocabulary). Finally, Chu and Schramm would expect performance on a pictorial test to be better if learning was accomplished through the use of visuals. On the other hand, the use of visuals might interfere with learning as measured by a test in print or audio form.

Radio and Tape Recordings

Radio. In reviewing instructional media research, Allen (1971) commented that by the end of World War II, instructional radio research in the United States "was virtually dead." Consistent with his observation, no experimental studies of the effectiveness of instructional radio were found to meet the criteria for this review, whether such investigations were conducted inside or outside the United States.

Chu and Schramm (1967) briefly noted radio research in 145 villages in Bombay State, India, which provided evidence of the superiority of a farm radio forum over the nonradio forum, based on pre- and posttests. They also cited a study by Cook in which pocket radios were used to broadcast Spanish drill exercises to college students enrolled in an elementary Spanish language course. Data indicated that the superiority found for the experimental over the control treatment was due to the greater amount of drill students got as a result of using the pocket radios.

Scupham (1970) summarised Danish research which assessed the comparative effectiveness of matched communications through print, radio, televised lecture and a more elaborate televised presentation using film, animation and other visual devices. The content of the instruction was the winter feeding of milch cows. Of the 500 subjects participating in the study, 120 were soldiers and the rest were students at agricultural colleges. No significant differences were found in the test scores of groups in the various treatment conditions. There were no significant differences between the two groups of soldiers who viewed the two television treatments, but there was a significant difference between the average score of the soldiers, whatever the presentation medium, and that of the students.

None of the above research was reported in sufficient detail to permit conclusions to be evaluated. The studies are mentioned merely to illustrate that radio has been used to present educational content under unusual and diverse conditions. The time is clearly ripe for some up-to-date and systematic research to be done to identify more precisely the conditions under which radio will have maximum instructional effectiveness, and the specific characteristics of radio that will bring about learning.

Tape recordings. Reviews of research on the use of tape recorders in language laboratories to present foreign language sounds indicated that there was very little experimental evidence as to the value of such recordings, especially at the post-secondary level (Chu and Schramm, 1967; McKeachie, 1967; Torkelson and Driscoll, 1968). Some individual studies of the use of tapes in language labs were found, but they failed to meet the criteria for this review.

The instructional effectiveness of tape-recorded instruction outside the language laboratory seems to have received even less attention. At

least no recent reviews of research could be located. Two experimental studies were found which appeared to meet the criteria for this review, although one exhibited certain design weaknesses which will be pointed out.

Stuck and Manatt (1970) conducted an investigation of the relative effectiveness of audio-tutorial and lecture methods of teaching concepts of school law to preservice teachers in the senior class at a state university in Iowa. The entire class of 219 students was randomly divided into two groups. The experimental students had one week in which to complete assignments covering 18 concepts of school law. The assignments, presented to the students individually by means of tape recordings in audio-tutorial booths, consisted of simulated situations which required the student to role-play a high school principal solving problems presented on the tape and to use available reference material in the process of identifying relevant law concepts. The control group attended seven hours of live lectures covering the same 18 law concepts. An outside panel of experts developed a criterion instrument based on these concepts. Before the test was used in the study, it was submitted to trial and revision, based on discrimination, difficulty and reliability analyses. Pre-posttest gain scores indicated that the audio-tutorial group was significantly superior to the traditionally taught group. It was determined statistically that neither prior teaching experience nor achievement level of the students was responsible for these results. The authors therefore concluded that the audio-tutorial technique was superior in teaching this subject matter.

Menne, Hannum, Klingensmith and Nord (1969) found no significant differences between taped and live lectures in an introductory college psychology course. Their experimental and control groups were formed by allowing students to choose between taking the course by taped or conventional lecture. In the experimental group (N=149) each student was issued a tape recorder, a complete set of lecture tapes, a booklet containing "the usual blackboard material" and a schedule of lecture topics to be given to the control group. Students also agreed not to attend live lectures, share tapes with control students or make copies of the tapes. They were allowed to proceed at their own pace and only attended classes when supplementary movies were shown. Students in the control group (N=211) attended the usual live lectures for the duration of the quarter. Both experimental and control groups took three, identical, objective, multiple-choice exams over the lecture material during the quarter. Data were submitted to covariance analysis "because it was recognized that experimental and control students could differ systematically." Given the voluntary basis on which students were assigned to groups, this expectation was certainly justified. It could be argued that the investigators improperly applied covariance analysis to their data. Use of this statistical technique is a dubious procedure at best when there are radical differences among groups. The investigators did, however, offer some assurance that the groups were similar on a number of cognitive variables. At any rate, no significant differences were found between the two groups on the basis of the objective course exam data. The investigators mentioned that attrition in the experimental group was practically nil (8 out of 149), but was over 10 percent in the control group. They speculated that in traditionally taught groups, dropouts were typically those who had fallen behind badly and saw no hope of recouping. For whatever reason, the very different attrition rates might indicate that selective factors were operating when one-tenth of the control students dropped the

course. The study was replicated during another quarter of the same academic year and similar, inconclusive results were obtained. In this study, too, the experimental students were all volunteers. One wonders how students would have performed had there been no alternative to taking the course by taped lectures. Moreover, what undetected biases resulted from the voluntary basis on which students were assigned to groups?

Multiple Media

A relatively new line of media research involves the use of several presentation devices in combination as an experimental treatment. Ideally, experimental designs should yield evidence on the extent to which each medium contributes to meeting instructional goals. It would then be appropriate to investigate the characteristics of a medium which make it particularly effective for accomplishing a given instructional objective. At present, however, most multi-media studies assess gross effects that the combined media produce on learning and do not provide for measuring the relative effectiveness of individual media.

Only two studies were found which seemed to meet the criteria for this review, both conducted by the same investigators (Edwards, Williams and Roderick, 1968). The first study compared the performance of college students learning first-semester business machines skills in an audiovisual-tutorial laboratory with the performance of a control group taught in the traditional manner. Assignment to either experimental or control group was by lottery. The control class was taught by the same teacher who prepared the materials for the experimental group, and both groups took identical final examinations at the end of the term. During the term, experimental students attended an open lab at any time convenient to them and received their instruction in business machines skills through programmed materials presented by continuous-loop sound film, and slides with tapes. These media were housed in individual carrels. As measured by the final performance test, the experimental group learned significantly more than the control group. A difference of even greater magnitude, also statistically significant, was noted when the scores of those students who had no previous exposure to business machines were analysed separately.

The same investigators conducted a separate but similar experiment in a beginning typing course. Students in the comparison group were taught to type by a teacher who used traditional methods in presenting the complete course in beginning typing. Students in the experimental group covered the same material in the audiovisual-tutorial lab situation, during which they were exposed to films, slides with tapes, tapes alone and printed instructions and tests. Students sat in individual learning carrel units especially designed to house the audiovisual devices. These devices were used for particular types of learning tasks. For example, continuous-loop film cartridges with synchronised narration were used for units in which demonstrations were necessary. Format and rules for typing manuscripts, cards and letters were prepared on slides accompanied by sound tapes. Tape recordings provided skill-building drills. Timed production and straight-copy drills were prepared on tapes which were used by students at practice stations with portable cassette tape recorders, rather than in carrels. Results on an end-of-term performance exam consisting of three-minute timed writings indicated that

students without prior typing training who learned beginning typing in the experimental situation significantly outperformed the traditionally trained group.

Limitations of Current Research and Suggestions for Future Research

The purpose and scope of this literature review were elaborated in the Introduction. In brief, an extensive literature search was made for research evidence which was relevant to selecting appropriate media for specified learning tasks. In particular, it was hoped that results of studies on the instructional effectiveness of media under a variety of learner and treatment conditions could be applied to the task of attempting to construct a media taxonomy. The disappointing result of the literature search was that little more than a dozen experimental studies were found to meet criteria which gave some assurance that findings were interpretable.

What is most impressive about the formidable body of literature surveyed for this review is that it shows that instructional media are being used extensively, under many diverse conditions, and that enormous amounts of money are being spent for the installation of very expensive equipment. All indications are that decisions as to which audiovisual devices to purchase, install and use have been based on administrative and organisational requirements and on considerations of cost, availability and user preference, not on evidence of instructional effectiveness. . . and no wonder. To date, media research in post-secondary education has not provided decision makers with practical, valid and dependable guidelines for making these choices on the basis of instructional effectiveness.

The aim of this section of the present discussion, then, is threefold:

1. To point out a few of the more common methodological weaknesses in current media studies.

2. To indicate that most media studies have left important questions unanswered, a shortcoming which makes the current literature of marginal use for the Council of Europe project.

3. To suggest types of media research which are needed to supply answers to these unanswered questions.

Common Methodological Defects in Media Research

Without going into a very lengthy and technical analysis, it is possible to enumerate briefly a few of the more common methodological weaknesses of recent media research in post-secondary education. Many of the items in the list were mentioned in the selection criteria outlined in the Introduction. Other items relate to basic tenets of sound experimental design. For a full discussion of these and other factors which can invalidate experimental findings, the reader is referred to Campbell and Stanley (1966) or to other standard source materials on experimental design (American Institutes for Research, 1970; Cox, 1958; Festinger and Katz, 1953; Rummel, 1964; Suchman, 1967).

The media studies which were not cited here, but which were surveyed in the literature search for this review, exhibited the following methodological defects pertaining to samples, treatments and measures of effectiveness.

Samples:

Failure to use random assignment methods.
Failure to report biases introduced by attrition of subjects during the course of the study.
Failure to control for pretreatment differences between groups.

Treatment:

Use of very brief exposures to treatments.
Failure to control for differences in exposure time for treatments being compared.
Use of some on-the-shelf medium that was not tailored to instructional objectives.
Description of task too brief to show what generalisations might be justified.
No evidence that media being compared gave equal coverage and emphasis to content.
When media were used in combination with some unspecified amount of teacher instruction, no analysis of relative contributions of each to learning outcomes.
Failure to control for teacher variable.

Measures of Effectiveness:

Use of tests with no evidence of reliability or validity.
Heavy reliance on subjective and affective measures of media effectiveness (judgments, unstructured observations, attitudes, preferences and testimonials).

In addition to all of the above shortcomings, many studies used criterion instruments which did not seem to be powerful enough to detect real differences between treatment conditions. Similarly, it often appeared that instructional materials were produced with insufficient skill to permit a fair test of teaching methods.

Questions Left Unanswered by Recent Media Research

The teacher making a choice among available audiovisual devices and materials on some logical basis must ask these sorts of questions:

Can this laboratory technique be taught best if my class sees a closeup, televised presentation of it?

What medium should I use to teach my class of illiterate farm workers how to speak English as a second language with good fluency, pronunciation and intonation?

Which of these lesson objectives would programmed instruction teach best?

When a teacher asks himself questions like these, he is concerned with media, learner and task variables that affect learning. Translated into the more precise language of educational research, questions like the following are relevant to making media choices:

What particular medium or combination of media will produce the most learning for Task A under Condition X, and why?

What factors in Condition X maximise learning of Task A by a particular medium or combination of media, and what factors interfere with such learning? Why?

What learner variables should be taken into account in deciding which students will profit most from instruction by a particular medium (or media) for Task A under Condition X?

If decisions about media selection and use are to be justified on the basis of instructional effectiveness, these are just the sorts of questions which need to be answered with valid and reliable experimental evidence.

Suggestions for Future Research

To attempt to base media choices on sound experimental evidence seems eminently sensible. The position taken here is that past media research has failed to provide the needed evidence; but that is not to say that it cannot be obtained. Future efforts, however, will have to use research strategies which are in marked contrast to those employed in most of the media studies surveyed for this review. These strategies relate to the planning of good research.

The question of which media to compare, or which learner and media characteristics to examine, should be determined in the light of subject-matter and task characteristics. At present, an entire unit or course is programmed, or produced as a series of televised lessons, or filmed, or tape-recorded or produced in multi-media format, without identifying specific instructional objectives to be met and without analysing the types and conditions of learning required. Learners are assigned to these experimental treatments without regard for traits which might interact with media and task characteristics.

The most illuminating evidence would come from research in which appropriate learner, media, task and situational variables are specified for use in multivariate designs (Allen, 1969, 1970; Campeau, 1966; Cronbach and Snow, 1969; Greenhill, 1967; Lesser and Schueler, 1966; McKeachie, 1967; Salomon and Snow, 1968, 1970; Siegel and Siegel, 1966; Snow, 1970; Snow and Salomon, 1968). In this way, not only would main treatment effects be detected, but interactions among variables might be found which could be used to enhance learning. This approach to media research assumes that potentially interacting variables can be defined, preferably according to some theoretical framework. The complexity of this task is evidenced by several recent attempts to identify variables for use in interactional research. Some of the more significant of these formulations may be briefly noted as alternatives to consider in planning future media research.

Siegel and Siegel (1966) reported a particularly painstaking and adept attempt to design and apply a multivariate paradigm in which learner, environmental, instructor and subject matter variables were taken into account. Several investigators have reported some ingenious methods for detecting individual differences which could form bases for assigning learners to treatment conditions (Cronbach and Snow, 1969; Ives, 1971a, 1971b; Snow, 1970; Snow and Salomon, 1968; Tallmadge and Shearer, 1969). Other investigators have outlined procedures for analysing subject matter and instructional objectives which comprise the experimental task (Briggs, 1970; Briggs and others, 1966; Gagne, 1965, 1968). Likewise, various approaches which might be used to specify media variables have also been suggested (Allen, 1970; Briggs, 1970; Salomon, 1970; Salomon and Snow, 1968; Tosti and Ball, 1969).

As Allen (1971) has observed, efforts like the above are being conducted within a theoretical framework, thus laying a foundation for a theory of instructional media. It would seem that if similar efforts were made in planning future media studies, experimental results would be more useful and generalisable than those produced by past media research. Further, systematic analyses of the media, learner, subject matter and situational factors involved in treatments being compared would permit more precise descriptions of such factors when the research is reported. Their description is quite inadequate in most of the current experimental literature.

In concluding this chapter, two further suggestions are offered to media researchers. The first calls for more multi-media studies based on rigorous analysis of the design elements in the instructional systems to be compared. That is, analysis of media, learner, subject matter and situational factors ought to form the basis for selecting media combinations which comprise the experimental treatments. The important point is that if multi-media studies are conducted within some theoretical framework, and if these studies meet the requirements of sound experimental design, more solid evidence will be obtained as to the manner in which several media can be combined to produce the maximum impact upon learning. This evidence cannot be provided directly by research which simply pits one medium against another medium. In addition, evidence derived from "basic" multi-media research would make a useful contribution to a much-needed theory of instructional media--one which takes into account the complexity of the learning process.

A final suggestion is that more effort be devoted to replication and follow-on studies. Almost no instances of either type of study were found in the experimental literature pertaining to the use of instructional media in post-secondary education. Replications are needed to confirm findings of a particular study or to show that those findings were merely artifacts of the unspecified variables which characterised the original study. Follow-on studies are needed to determine under what conditions, for which students and for which learning tasks particular instructional media will produce the most learning.

Summary

The purpose of this literature review was to summarise results of experimental studies on the instructional effectiveness of audiovisual media

FILMED FROM BEST AVAILABLE COPY

in post-secondary education. To ensure that a study's findings were interpretable, seven major screening criteria were employed to gauge the nature and quality of published evaluation data and the soundness of the study's experimental design. In general, a study was accepted for review if media effectiveness was evaluated by comparing the performance of experimental and control groups on objective measures of cognitive achievement. Evaluations based on noncognitive data were not considered, e.g., attitudinal, motivational and media-preference studies. Also omitted from review were assessments of media effectiveness derived from informal evaluations, questionnaire and opinionnaire responses, enrollment and completion statistics, or surveys of costs, services and use of audiovisual facilities.

Computerised and manual search procedures identified over 1200 research and research-related references, but only 200 of the more promising items were studied in any detail. Of these, little more than a dozen studies were found to meet the screening criteria, and most were programmed instruction studies. Reviews of research and findings from individual studies were summarised under the following headings: programmed instruction; television, videotape and motion pictures; slides, filmstrips, overhead transparencies and still pictures; radio and tape recordings; and multiple media.

A closing section enumerated particularly common methodological defects in recent media research and suggested new strategies for future media studies. More effort should be directed toward planning good research and to posing appropriate research questions. Recent formulations for defining learner, task, media and situational variables were cited. It was proposed that potentially interacting variables be (1) defined according to some theoretical framework, (2) used in multivariate designs and (3) taken into account in planning more multi-media studies. In addition to improving the theoretical and methodological quality of future media research, it was also proposed that more effort be devoted to replication and follow-on studies. The purpose of such studies would be to confirm and extend our understanding of those conditions under which particular media will have maximum instructional effectiveness.

References

Note: Documents for which an ED number is given are available from the ERIC Document Reproduction Service in two forms: microfiche (4- by 6-inch sheets of microfilm on which up to 70 pages of text can be reproduced) or hard copy (reproductions of the document on paper in its original size). Purchase prices, not given below, will be provided upon request. Write to ERIC Document Reproduction Service, P.O. Drawer O, Bethesda, Maryland, 20014, U.S.A. It is vital that the document be identified by its ED number.

Alexander, J. E. Vocabulary improvement methods, college level. Knoxville, Tennessee: Tennessee University, 1970. ED 039 095

Allen, W. H. Technology: Roles for men and machines. In D. G. Hays (Ed.), Britannica review of American education. Vol. 1. Chicago, Illinois: Encyclopaedia Britannica, 1969. Pp. 139-157.

Allen, W. H. Categories of instructional media research. In G. Salomon and R. E. Snow (Eds.), Commentaries on research in instructional media: An examination of conceptual schemes. Viewpoints (Bulletin of the School of Education, Indiana University), 1970, 46 (5), 1-14. ED 044 900

Allen, W. H. Instructional media research. Past, present, future. Audiovisual Communication Review, 1971, 19, 5-18.

American Institutes for Research. Evaluative research; strategies and methods. Pittsburgh, Pennsylvania: AIR, 1970.

Anderson, R. C. Educational psychology. Annual Review of Psychology, 1967, 18, 129-164.

Barnes, M. R. An experimental study of the use of programmed instruction in a university physical science laboratory. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Minneapolis, Minnesota, 1970. ED 041 726

Briggs, L. J. Learner variables and educational media. Review of Educational Research, 1968, 38, 160-176.

Briggs, L. J. Handbook of procedures for the design of instruction. Pittsburgh, Pennsylvania: American Institutes for Research, 1970. Monograph No. 4.

Briggs, L. J., Campeau, P. L., Gagné, R. M. and May, M. A. Instructional media: A procedure for the design of multi-media instruction, a critical review of research and suggestions for future research. Pittsburgh, Pennsylvania: American Institutes for Research, 1966. Monograph No. 2.

Campbell, D. T. and Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago, Illinois: Rand-McNally, 1966.

Campeau, P. L. Selective review of literature on audiovisual media of instruction. In L. J. Briggs and others, Instructional media: A procedure for the design of multi-media instruction (abbreviated title). Pittsburgh, Pennsylvania: American Institutes for Research, 1966. Pp. 99-142. Monograph No. 2.

Chu, G. C. and Schramm, W. Learning from television: What the research says. Stanford, California: Stanford University, Institute for Communication Research, 1967. ED 014 900

Coppen, H. (Ed.) A survey of British research in audio-visual aids, 1945-1967. London: National Committee for Audio-Visual Aids in Education, Educational Foundation for Visual Aids, 1968.

Coppen, H. (Ed.) A survey of British research in audio-visual aids. Supplement 1, 1968. London: National Committee for Audio-Visual Aids in Education, Educational Foundation for Visual Aids, 1969.

Coppen, H. (Ed.) A survey of British research in audio-visual aids. Supplement 2, 1969. London: National Committee for Audio-Visual Aids in Education, Educational Foundation for Visual Aids, 1970.

Coppen, H. (Ed.) A survey of British research in audio-visual aids. Supplement 3, 1970. London: National Committee for Audio-Visual Aids in Education, Educational Foundation for Visual Aids, 1971.

Cox, D. R. Planning of experiments. New York: John Wiley & Sons, 1958.

Cronbach, L. J. and Snow, R. E. Individual differences in learning ability as a function of instructional variables: Final report. Stanford, California: Stanford University, School of Education, 1969.

Daniel, W. J. and Murdoch, P. Effectiveness of learning from a programmed text compared with a conventional text covering the same material. Journal of Educational Psychology, 1968, 59, 425-451.

Dubin, R., Hedley, R. A., (with) Schmidbauer, M., Goldman, D. R. and Traveggia, T. C. The medium may be related to the message: College instruction by TV. Eugene, Oregon: University of Oregon, Center for the Advanced Study of Educational Administration, 1969. ED 031 926

Edwards, R. K., Williams, M. L. and Roderick, W. W. An experimental pilot study to explore the use of an audio-visual-tutorial laboratory in the secretarial offerings at the community college level in Michigan. Lansing, Michigan: Lansing Community College, 1968. ED 027 387

Festinger, L. and Katz, D. (Eds.) Research methods in the behavioral sciences. New York: Holt, Rinehart & Winston, 1953.

Frase, L. T. Boundary conditions for mathemagenic behaviors. Review of Educational Research, 1970, 40, 337-347.

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart & Winston, 1965.

Gagné, R. M. Media and the learning process. Paper presented to the DAVI Conference, Houston, Texas, March 25, 1968. ED 022 368

Gagné, R. M. and Rohwer, W. D., Jr. Instructional psychology. Annual Review of Psychology, 1969, 20, 381-418.

Giese, D. L. and Stockdale, W. A. Comparing an experimental and a conventional method of teaching linguistic skills. The General College Studies, 1966, 2 (3), 1-10. ED 019 072

Greenhill, L. P. Review of trends in research on instructional television and film. (Introduction to) J. C. Reid and D. W. MacLennan, Research in instructional television and film. Washington, D. C.: U.S. Government Printing Office, 1967. Pp. 1-17. ED 015 673

Hamilton, R. S. and Heinkel, O. A. English A: An evaluation of programmed instruction. San Diego, California: San Diego City College, 1967. ED 013 619

Hartley, J. R. An experiment showing some student benefits against behavioural costs in using programmed instruction. Programmed Learning and Educational Technology, 1968, 5, 219-229.

Ives, J. M. A strategy for instructional television research. Audio-visual Communication Review, 1971, 19, 149-160. (a)

Ives, J. M. Research in educational television. Programmed Learning and Educational Technology, 1971, 8, 161-172. (b)

Kanner, J. H. The instructional effectiveness of color in television: A review of the evidence. Stanford, California: Stanford University, ERIC Clearinghouse on Educational Media and Technology, 1968. ED 015 675

Karis, C., Kent, A. and Gilbert, J. E. The interactive effect of responses per frame, response mode, and response confirmation on intraframe S-R association strength. Final report. Boston, Massachusetts: Northeastern University, 1970. ED 040 591

Lesser, G. S. and Schueler, H. New media research in teacher education. Audiovisual Communication Review, 1966, 14, 318-361.

Marsh, L. A. and Pierce-Jones, J. Programmed instruction as an adjunct to a course in adolescent psychology. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois, February 9, 1968. ED 018 989

McKeachie, W. J. New developments in teaching: New dimensions in higher education. No. 16. Durham, North Carolina: Duke University, 1967. ED 013 341

Menne, J. W., Hannum, T. E., Klingensmith, J. E. and Nord, D. Use of taped lectures to replace class attendance. Audiovisual Communication Review, 1969, 17, 42-46.

Reid, J. C. and MacLennan, D. W. Research in instructional television and film. Washington, D. C.: U.S. Government Printing Office, 1967. ED 015 673

Rich, O. S., Poll, R. D. and Williams, T. M. The utilization of large-screen TV to overcome shortages of classroom space and teaching personnel. Provo, Utah: Brigham Young University, 1966. ED 010 175

Rothkopf, E. Z. The concept of mathemagenic activities. Review of Educational Research, 1970, 40, 325-336.

Rummel, J. F. An introduction to research procedures in education. (2nd ed.) New York: Harper & Row, 1964.

Salomon, G. What does it do to Johnny? A cognitive-functionalistic view of research on media. In G. Salomon and R. E. Snow (Eds.), Commentaries on research in instructional media: An examination of conceptual schemes. Viewpoints (Bulletin of the School of Education, Indiana University), 1970, 46 (5), 33-62. ED 044 900

Salomon, G. and Snow, R. E. The specification of film attributes for psychological and educational research purposes. Audiovisual Communication Review, 1968, 16, 225-244.

Salomon G. and Snow, R. E. (Eds.) Commentaries on research in instructional media: An examination of conceptual schemes. Viewpoints (Bulletin of the School of Education, Indiana University), 1970, 46 (5). Entire issue devoted to theoretical and methodological issues pertaining to media research. ED 044 900

Scupham, J. Broadcasting and the Open University. Journal of Educational Technology, 1970, 1, 44-51.

Siegel, L. and Siegel, L. C. The instructional gestalt in televised university courses. Oxford, Ohio: Miami University, 1966. ED 010 177

Smith, I. M. A further experimental study of the effect of television broadcasts on the G courses in engineering science: Part 1. The Vocational Aspect, 1968, 20 (45), 78-85. (a)

Smith, I. M. An experimental study of the effect of television broadcasts on the G courses in engineering science: Part 2. The Vocational Aspect, 1968, 20 (46), 89-100. (b)

Snow, R. E. Research on media and aptitudes. In G. Salomon and R. E. Snow (Eds.), Commentaries on research in instructional media: An examination of conceptual schemes. Viewpoints (Bulletin of the School of Education, Indiana University), 1970, 46 (5), 63-89. ED 044 900

Snow, R. E. and Salomon, G. Aptitudes and instructional media. Audiovisual Communication Review, 1968, 16, 341-357.

Stuck, D. L. and Manatt, R. P. A comparison of audio-tutorial and lecture methods of teaching. Journal of Educational Research, 1970, 63, 414-418.

Suchman, E. A. Evaluative research: Principles and practice in public service and social action programs. New York: Russel Sage Foundation, 1967.

Tallmadge, G. K. and Shearer, J. S. Relationships among learning styles, instructional methods, and the nature of learning experiences. Journal of Educational Psychology, 1969, 60, 222-230.

Tobias, S. The effect of creativity, response mode, and subject matter familiarity on achievement from programmed instruction. New York: City University of New York, 1968. ED 027 756

Tobias, S. The effect of creativity, response mode, and subject matter familiarity on achievement from programmed instruction. Journal of Educational Psychology, 1969, 60, 453-460.

Tobias, S. and Abramson, T. The relationship of anxiety, response mode, and content difficulty to achievement in programed instruction. Final report. New York: City University of New York, New York City College, 1970. ED 040 605

Torkelson, G. M. and Driscoll, J. P. Utilization and management of learning resources. Review of Educational Research, 1968, 38, 129-159.

Tosti, D. T. and Ball, J. R. A behavioral approach to instructional design and media selection. Audiovisual Communication Review, 1969, 17, 5-25.

Twyford, L. C., Jr. Educational communications media. In R. L. Ebel (Ed.), Encyclopedia of Educational Research (4th ed.) London: Macmillan, Collier-Macmillan, Ltd., 1969. Pp. 367-380.

Unwin, D. An organisational explanation for certain retention and correlation factors in a comparison between two teaching methods. Programmed Learning (and Educational Technology), 1966, 3, 35-39.

Valverde, H. and Morgan, R. L. Influence on student achievement of redundancy in self-instructional materials. Programmed Learning and Educational Technology, 1970, 7, 194-199.

Wilds, P. L. and Zachert, V. Effectiveness of a programed text in teaching gynecologic oncology to junior medical students, a source book on the development of programed materials for use in a clinical discipline. Augusta, Georgia: Medical College of Georgia, 1966. ED 010 024